Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **514,** 4239–4245 (2022)
Advance Access publication 2022 June 14

https://doi.org/10.1093/mnras/stac1597

Identification of new hot subdwarf binary systems by means of Virtual Observatory tools

E. Solano,[1]★ A. Ulla,[2] E. Pérez-Fernández,[2,3] C. Rodrigo,[1] R. Oreiro,[2,4] A. Aller ,[1] M. Manteiga,[5] R. Santoveña-Gómez,[6] M. A. Álvarez [6] and C. Dafonte[6]

[1]*Departmento de Astrofísica, Centro de Astrobiología (CSIC-INTA), ESAC Campus, Camino Bajo del Castillo s/n, E-28692 Villanueva de la Cañada, Madrid, Spain*
[2]*Applied Physics Department, Universidade de Vigo (UVIGO), Campus Lagoas-Marcosende, s/n, E-36310 Vigo, Spain*
[3]*IES de Beade, Consellería de Educación e Ordenación Universitaria, Camino do Outeiro 10, E-36312 Vigo, Spain*
[4]*IES Ramón Llull, Carrer de Ramon Llull, 10, E-46021 València, Spain*
[5]*CIGUS CITIC - Department of Nautical Sciences and Marine Engineering, University of A Coruña, Paseo de Ronda 51, E-15011 A Coruña, Spain*
[6]*CIGUS CITIC - Department of Computer Science and Information Technologies, University of A Coruñan, Campus Elviña s/n, E-15071 A Coruña, Spain*

Accepted 2022 June 2. Received 2022 May 24; in original form 2022 March 30

ABSTRACT

The estimation of the binary fraction of hot subdwarfs is key to shed light on the different evolution scenarios proposed to explain the loss of the hydrogen envelope during the red giant branch phase. In this paper, we analyse the spectral energy distribution of the hot subdwarfs included in a recent and comprehensive catalogue with the aim of identifying companions. Our methodology shows a performance superior to the photometric criteria used in that study, identifying 202 objects wrongly classified as binaries according to their spectral energy distributions, and finding 269 new binaries. Out of an initial sample of 3186 objects, we classified 2469 as single and 615 as binary hot subdwarfs. The rest of the objects (102) were not classified because of their inadequate spectral energy distribution fitting due, in turn, to poor quality photometry. Effective temperatures, luminosities, and radii were computed for 192 singles and 42 binaries. They, in particular the binary sample, constitute an excellent data set to further perform a more careful spectroscopic analysis that could provide detailed values for the chemical composition, masses, ages, rotation properties, or reflection effects for the shortest period systems. The results obtained in this paper will be used as a reference for a forthcoming work where we aim to generalize binary and single hot subdwarf classification using Artificial Intelligence-based techniques.

Key words: astronomical data bases: miscellaneous – virtual observatory tools.

1 INTRODUCTION

Hot subdwarf (hot sd) stars are core-helium burning stars at the blue end of the horizontal branch or even beyond that stage, in the region known as the extreme horizontal branch. They are often located at high galactic latitudes and, according to atmospheric composition, they are mainly divided into B- (sdB) or O- (sdO) types, if hydrogen or helium dominated, respectively (Drilling et al. 2013).

With effective temperatures exceeding 19 000 K, $\log g \geq 4.5$ dex, radii of a few tenths of a solar radius and masses of ~ 0.5 $M_\odot$, hot sds are objects that have lost most of their hydrogen envelope ($M_{\rm env} < 0.01$ $M_\odot$) during the red giant branch phase (Heber 2009). Therefore, they are unable to sustain hydrogen shell burning and thus, to follow canonical evolution through the asymptotic giant branch proceeding, instead, directly towards the white dwarf cooling track. Circumstances that lead to the removal of all but a tiny fraction of the hydrogen envelope are still a matter of debate. Theoretical evolution scenarios proposed so far include enhancement of the mass-loss efficiency near the red giant branch tip (D'Cruz et al. 1996) or mass-loss, mass-transfer, or coalescence through binary interaction (Pelisoli et al. 2020). Further and comprehensive information on the hot sd properties can be found in Heber (2016).

Since the first large area surveys made in the second half of the 20th century (e.g. Kilkenny, Heber & Drilling 1988), the number of known hot sds has dramatically grown. In part, this is thanks to the advent of large spectroscopic surveys like MUCHFUSS (Geier et al. 2011), SDSS (Geier et al. 2015; Kepler et al. 2015, 2016), or LAMOST (Lei et al. 2019). Geier (2020) includes a recent and up-to-date compilation of hot sd catalogues.

In parallel to this data increase, the advent of initiatives like the Virtual Observatory (VO)[1] has opened new lines of work in the field of massive data analysis. Oreiro et al. (2011), using a VO methodology, defined a procedure particularly designed to uncover uncatalogued hot sds among blue object samples, with a low contamination factor over white dwarfs, cataclysmic variables, and OB stars. Out of 38 new candidates, 30 of them with spectra available, they could confirm 26 new hot sds and reported, at least, eight binary systems. The methodology described in that paper was further extended by Pérez-Fernández et al. (2016) who reported 65 new hot sds out of 68 candidates with SDSS spectrum.

★ E-mail: esm@cab.inta-csic.es

[1]http://www.ivoa.net

It is well known that a significant fraction of hot sds resides in binary systems. Studies based on optical colours, radial velocities, high-resolution optical spectroscopy or excess in the near-infrared have led to a fraction of hot sd binaries ranging from 20 to 66 per cent depending on the methodology, inhomogeneity of data sets, or selection effects (Girven et al. 2012).

In this paper, we will test the ability of VO tools to uncover new hot sds in binary pairs, and/or to confirm proposed binary candidates using the Geier (2020) catalogue of known hot sd stars. The article is organized as follows: after the Introduction (Section 1), we describe in Section 2 the object selection, which is based on good determinations of *Gaia* Early Data Release 3 (EDR3) parallaxes and colours and on the expected locus of hot sds in a colour–absolute magnitude diagram (Section 2.1), the procedure followed to build the spectral energy distributions (SEDs; Section 2.2), and the identification of binaries (Section 2.3). Section 3 presents our determinations of effective temperature for single and binary targets. Finally, in Section 4, we summarize the results of this work and give indications on how to proceed in future binary searches on larger hot sds catalogues.

2 ANALYSIS

2.1 Object selection

The catalogue of hot sds compiled by Geier (2020) contains 5874 unique sources (2187 with spectroscopically derived physical parameters taken from the literature), out of which 528 are reported as new hot sds. Also, 777 and 1365 sources were classified as binaries depending on whether a spectroscopic or a photometric approach was used, respectively. We use this catalogue as a reference throughout this work.

We perform a first filtering of the catalogue keeping only those objects with low relative errors in magnitudes and parallaxes. We are interested in keeping sources with good values of distances as they will be used as filtering criterion in the determination of effective temperatures (Section 3). With this purpose, we first cross-matched the 5874 objects of Geier (2020) with the *Gaia* EDR3 catalogue (Gaia Collaboration 2020) using a 5 arcsec radius. We found 5862 sources with counterparts in *Gaia* EDR3. Since the Geier's catalogue avoids the Galactic plane, the contamination from matching to other stars within this search radius can be considered as negligible. Anyway, in those cases where more than one counterpart exists in the search region, only the nearest one to the Geier's source was considered. Then, we kept only counterparts with relative errors of less than 10 per cent in G, $G_{\rm BP}$, and $G_{\rm RP}$ and less than 20 per cent in parallax. The absolute *Gaia* magnitude in the G band was estimated using

$$M_{\rm G} = G + 5\log\varpi + 5, \quad (1)$$

where ϖ is the parallax in arcseconds. In our case, the inverse of the parallax is a reliable distance estimator because we kept only sources with relative errors in parallax lower than 20 per cent (Luri et al. 2018). After this filtering, 3571 sources were kept. The great majority of the sources (all but four) were rejected because they did not fulfill the condition on parallaxes. Then, we applied the colour and absolute magnitude selection criteria described in Geier et al. (2019) to define the expected locus for hot sds. In particular, we kept objects occupying the region $-0.7 < G_{\rm BP}–G_{\rm RP} \leq 0.7$ and $-1.0 < M_{\rm G} < 7.0$ mag and fulfilling equations (1)–(3) in that paper. This selection resulted in 3186 objects (Fig. 1).

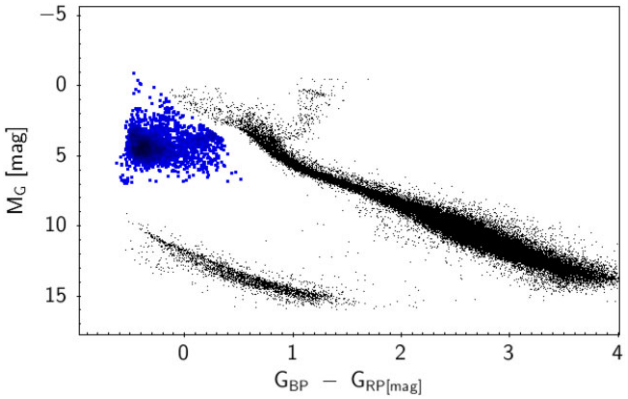

Figure 1. Colour–magnitude diagram built using *Gaia* EDR3 sources with parallaxes larger than 10 mas, photometric errors in G, $G_{\rm BP}$, and $G_{\rm RP}$ less than 10 per cent, RUWE < 1 to remove binaries (Arenou et al. 2018; Lindegren et al. 2018, 2021) and the colour excess factor applied following the approach described in Riello et al. (2021) (black dots). The 3186 hot sd stars selected following the criteria described in Section 2.1 are overplotted as blue dots.

2.2 SED building

For the 3186 objects selected in the previous section, we used VOSA[2] (Bayo et al. 2008) to build their SEDs. VOSA is a tool developed by the Spanish Virtual Observatory[3] designed to build the SEDs of thousands of objects at a time from a large number of photometric catalogues, ranging from the ultraviolet (UV) to the infrared. VOSA compares catalogue photometry with different collections of theoretical models and determines which model best reproduces the observed data following different statistical approaches. Physical parameters can then be estimated for each object from the model that best fits the data.

In our analysis, the photometric information was obtained from the following catalogues: GALEX (Bianchi & GALEX Team 2000), Pan-STARRS DR2 (Magnier et al. 2020), *Gaia* (Gaia Collaboration 2020), SDSS DR12 (Alam et al. 2015), 2MASS-PSC (Skrutskie et al. 2006), and WISE (Wright et al. 2010).

2.3 Binary identification

VOSA also allows the identification of flux excess in an SED due to the presence of, for instance, a disc, a nebula, or a close companion. To this end, VOSA first executes an iterative algorithm which is an extension of the method described in Lada et al. (2006). Starting at $\lambda \geq 21\,500$ Å, VOSA computes the slope of the linear regression of the observational SED in a $\log \nu F_\nu$ versus $\log \nu$ diagram. This slope is recomputed by adding new infrared photometric points at every step. If, in any of these steps, the slope becomes significantly smaller than the one expected from a stellar photospheric emission, VOSA flags the object as potentially affected by excess and photometric points at longer wavelengths are not taken into account in the SED fitting process.

Once the SED fitting is completed, VOSA performs a further refinement of the excess estimation by comparing, for each photometric point, the observational flux to the synthetic flux obtained from the model that best fits the data. Significant ($> 3\sigma$) deviations in the observational flux are flagged by VOSA as potential IR excesses. A

[2]http://svo2.cab.inta-csic.es/theory/vosa/
[3]http://svo.cab.inta-csic.es

detailed description of how VOSA manages the infrared excess can be found in the VOSA documentation.[4]

From the sample of 3186 objects, VOSA identified 2469 single objects and 615 binary systems. The rest of the objects (102) were not classified because of their bad SED fitting due to, for instance, poor quality photometry or lack of enough photometric points. This gives a binarity ratio of 20 per cent [615/(615 + 2469)], in agreement with Geier (2020) [13 per cent (777/5874) or 23 per cent (1365/5874), depending on if, in their study, spectroscopic of photometric binaries are considered]. We note that our estimation can be biased by selection effects indicating just a lower limit to the real number of binaries as our detection of companions is restricted to not too cool objects (that would be out-shined by the hot sd in the optical regime), not too hot objects (that would out-shine the hot sd) as well as to systems where the difference in effective temperature between the hot sd and the companion is significant (systems with similar $T_{\rm eff}$ will be seen as a single SED by VOSA). Also, objects without good photometric information in the 2MASS-WISE regime (which is where the flux excess indicating the presence of a companion occurs) could be wrongly classified as singles. In fact, 232 out of the 2469 objects classified as singles lie in this category. Assuming a binarity ratio of 20 per cent, this would imply that $\sim$ 50 sources classified as single could be real binaries. Moreover, hot sds showing large reflection effects will have associated SEDs with large photometric variability as the photometric data to construct the SED were obtained at different observing epochs. They will have a bad SED fitting lying, thus, among the 102 not classified objects. However, even in the unrealistic scenario in which the 102 objects with bad SED fitting are true reflection binaries, the binarity ratio will not change too much: from 20 per cent [615/(615+2469)] to 22.5 per cent [717/(717 + 2469)]. Finally, wide binaries escape from our search as they do not present a composite SED, although Pelisoli et al. (2020) suggested that only a small fraction of hot sds are in wide binaries.

Geier et al. (2017) provide a classification of hot sds in three classes (sdOs, sdBs, and sd + MS) following a purely empirical scheme based on photometry. Geier (2020) expands this methodology using SDSS (Alam et al. 2015), APASS (Henden et al. 2015), Pan-STARRS1 (Chambers et al. 2016), and SKyMapper (Wolf et al. 2018) data. The SDSS- and SkyMapper-based colour classes are regarded as the most trustworthy because the $u - g$ colour allows to distinguish between sdB and the hotter sdO types better than any other colour combination.

Among our list of 3084 objects (2469 singles + 615 binaries), there are 548 photometrically classified as binaries by Geier (2020) using any of their colour–colour combinations. However, we found that, for 202 of them (37 per cent), the SED can be better fitted (in terms of χ^2) using a single object model (see Fig. 2 as an example). Only 24 of them are spectroscopically confirmed as binaries in Geier (2020), which reinforces the hypothesis of their single nature and indicates that VOSA offers a superior performance discriminating between singles and binaries than the photometric criteria used in Geier (2020).

As shown in Fig. 3, the majority of these sources (141/202, 70 per cent) are located at $G_{\rm BP} - G_{\rm RP} \leq 0.0$, which is the locus for single hot sds according to Geier et al. (2019). If only SDSS photometry is used, then 253 sources are classified as binaries, out of which 66 (26 per cent) are classified as single by VOSA. 54 of these 66 sources (89 per cent) have negative $G_{\rm BP} - G_{\rm RP}$ values. Besides, among the 615 binary systems identified by VOSA, 269 objects were not reported as binaries in Geier (2020).

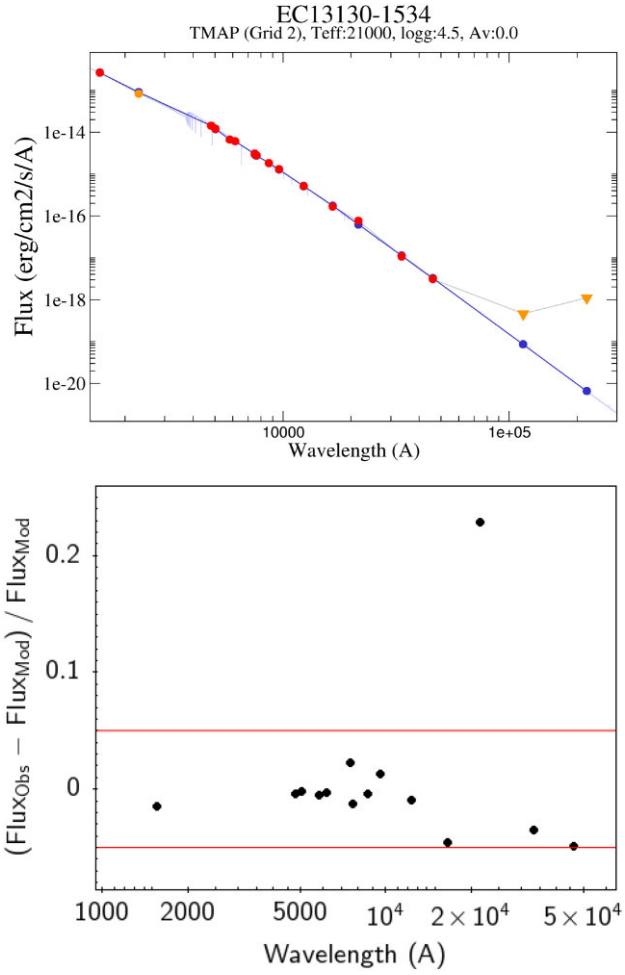

Figure 2. Top: SEDs of EC13130-1534, photometrically classified as binary in Geier (2020) but whose model fitting with VOSA indicates that it is most likely a single star. The red dots represent the observed photometry while the blue spectrum indicates the theoretical model that fits best. The yellow inverted triangle indicates that the photometric value corresponds to an upper limit and, thus, is not taken into account in the fitting process. Bottom: Residuals of the fit for the good photometric points. Except for K_s, the differences between the observed and the model flux are below 5 per cent (red lines).

3 EFFECTIVE TEMPERATURE DETERMINATION

3.1 Single SED objects

For the 2469 objects classified as single by VOSA, we attempted to estimate their effective temperatures. For this, we made use of the TMAP (Tubingen NLTE Model Atmosphere Package) collection of theoretical models (Rauch & Deetjen 2003). Hydrogen-only models in the 20 000 K $< T_{\rm eff} <$ 150 000 K range were requested to TheoSSA[5] and implemented in the theoretical spectra web server of Spanish Virtual Observatory portal[6] for their use by VOSA. In our analysis, the range in surface gravity was restricted to log g: 4.5–6 dex, typical of hot sds.

Extinction can play an important role in shaping the SED in particular at short wavelengths where the peak of the SED is reached.

[4] https://bit.ly/2KRCv9x

[5] http://dc.zah.uni-heidelberg.de/theossa/q/web/form

[6] http://svo2.cab.inta-csic.es/theory/newov2

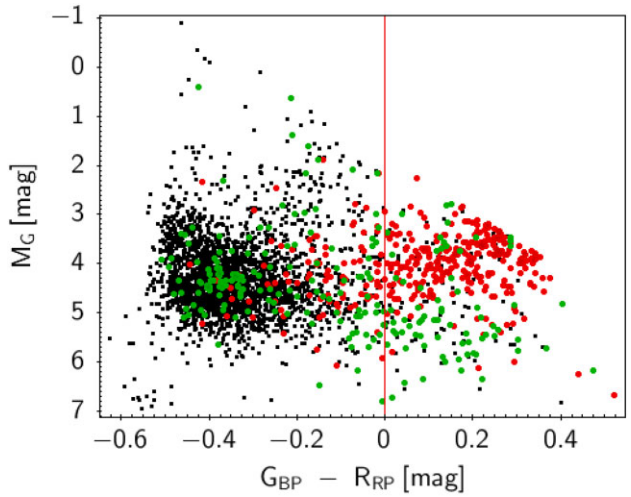

Figure 3. Colour–magnitude diagram showing the initial list of 3084 objects (black dots), the 548 objects classified as binaries by Geier (2020) (red dots), and the 202 objects classified as binaries by Geier (2020) but as single sources by VOSA (green dots). The vertical solid line represents the frontier between the binary (G_{BP} – G_{RP} with positive values) and single (G_{BP} – G_{RP} with negative values) locii as described in Geier et al. (2019).

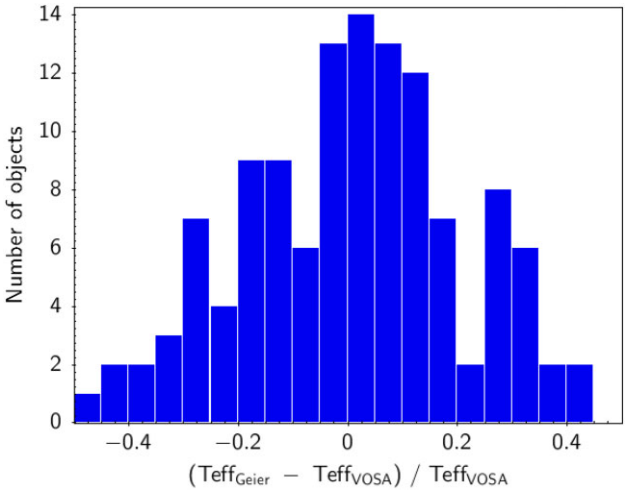

Figure 4. Relative differences between the spectroscopic temperatures given in Geier (2020) and those derived using VOSA for the 98 objects in common fulfilling the conditions described in Section 3.1. More than 70 per cent of the objects have relative errors below 20 per cent with VOSA temperatures being slightly lower than those given in Geier (2020). This could be associated with an underestimation of the extinction factor, Av.

If extinction is underestimated, the slope of the SED will appear flattened at short wavelengths and the derived effective temperature will be lower. To account for this effect and to minimize the extinction–effective temperature degeneracy in the SED fitting, we decided to leave extinction as a free parameter in the SED fitting process taking values in the range $0 \leq \mathrm{Av} \leq 1$ mag and keep only objects at distances < 1000 pc. VOSA estimates extinction using the laws described in Fitzpatrick (1999) (optical) and Indebetouw et al. (2005) (infrared). Distances were obtained from *Gaia* EDR3 in the way explained in Section 2.1.

After this, we kept only the objects fulfilling the following filtering criteria:

(i) Given the high temperatures of the hot sds, the accuracy of the values estimated from the SED fitting using VOSA is closely linked to the photometric coverage in the blue part of the electromagnetic spectrum. Therefore, we kept only objects having both GALEX FUV and NUV photometry. GALEX photometry was corrected using the expressions given in Camarota & Holberg (2014, table 5)

(ii) We selected objects with $T_{\rm eff} > 20\,000$ K, to avoid boundary issues as our grid of models starts at $T_{\rm eff} = 20\,000$ K.

(iii) We selected objects with $T_{\rm eff} < 50\,000$ K. Higher temperatures will be affected by large uncertainties due to the lack of photometric information at wavelengths shorter than GALEX. According to the Wien's law, the SED of an object with $T_{\rm eff} = 50\,000$ K peaks at 576 Å, well beyond the GALEX FUV band ($\lambda_{\rm mean} = 1546$ Å). Although this cut in effective temperature implies that hot sdOs will not be considered, this should not have a major impact on the final number of candidates since, as Heber (2016) pointed out, the large majority of hot sds have a temperature between 20 000 and 40 000 K.

(iv) Good SED fitting (*vgfb* < 15). *vgfb* is a modified χ^2, internal to VOSA, and calculated by forcing $\Delta F_{\rm bol}$ to be, at least, $0.1 \times F_{\rm bol}$, where $\Delta F_{\rm bol}$ is the error in the total observed bolometric flux. This parameter is particularly useful when the photometric errors of any of the catalogues used to build the SED are underestimated. *vgfb* $<$ 15 is a reliable indicator for a good fit.

After applying the previous conditions, only 192 sources were kept, out of which, 98 have determinations of effective temperature in Geier (2020). This subset of 98 objects was used to assess the reliability and robustness of the effective temperatures calculated with VOSA. Fig. 4 shows the relative differences between these values and those derived with VOSA. The agreement is good with VOSA temperatures being slightly lower than those in Geier (2020). Information on the 192 objects for which effective temperatures were estimated can be found at the SVO archive of hot sds (see Appendix).

3.2 Binary SED objects

If there is a difference in $T_{\rm eff}$ between the hot sd and the main-sequence companion (with the hot sd mostly contributing at shorter wavelengths while the companion dominates at longer wavelengths), VOSA is able to decompose the observed SED. Physical parameters can then be estimated for each one of the components by taking advantage of the two-body fitting algorithm implemented in VOSA (see the VOSA documentation[7] for a detailed description on how the two-body SED fitting is made).

To ensure the reliability of the results provided by our methodology, we first applied it to known composite hot sds with spectroscopically derived effective temperatures. In particular, we used the list of objects compiled in Girven et al. (2012) and Németh, Kawka & Vennes (2012), which provide effective temperatures for both components for 227 and 29 objects, respectively. After applying the restrictions in terms of distances (parallax >1 and relative error < 20 per cent), GALEX photometry and good SED fitting, we ended up with 6 and 11 objects, respectively. SEDs were fitted using the TMAP grid of models for the hot sd component and BT-Settl models (Allard, Homeier & Freytag 2012) for the cooler companion. The ranges of effective temperatures and surface gravities were set to 20 000–50 000 K and 4.5–6 dex and 2000–19 000 K and 4–6 dex, for the hot sd and the companion, respectively.

[7] http://svo2.cab.inta-csic.es/theory/vosa50/helpw4.php?otype=starnew&what=action=help&what = fitbin

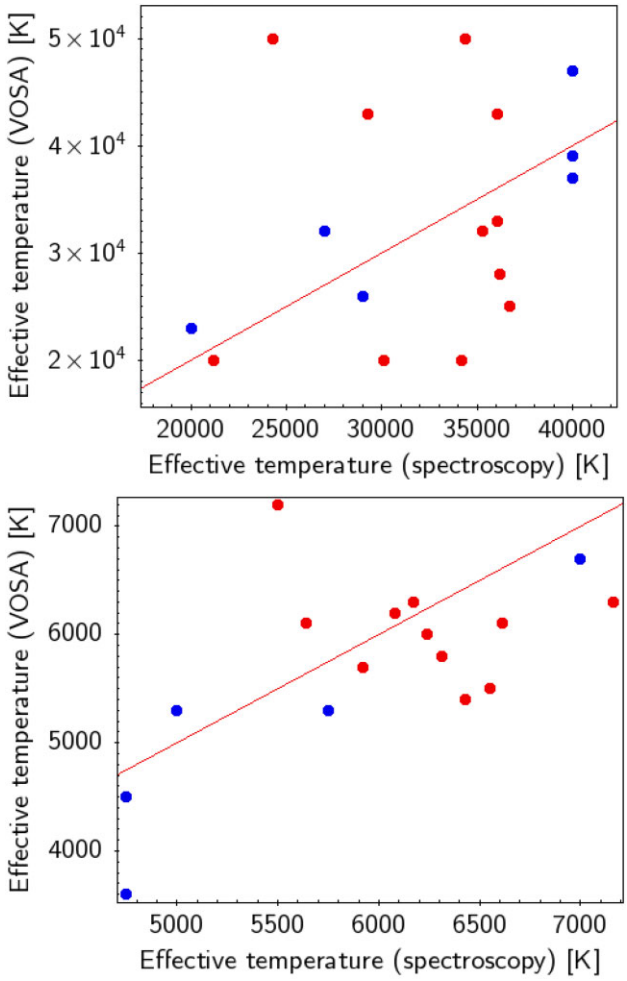

Figure 5. Comparison between the effective temperatures estimated with VOSA and the corresponding values taken from Girven et al. (2012) (blue) and Németh et al. (2012) (red) for the hot sds (top) and the cool companions (bottom).

In Fig. 5, we compare the effective temperatures provided by VOSA with the spectroscopic determinations for those 17 systems. We find that there are no systematic differences. No companions cooler than K-types are detected because they are completely outshined by the hot sd. Similarly, companions hotter than A-types would dominate the SED flux, making the hot sd unseen. The spread in temperatures, in particular for the hot sds, can be ascribed to the following reasons:

(i) As previously mentioned, extinction is left as a free parameter in the VOSA two-body fitting process. A_V ranges from 0.0 to 1.0 mag in steps of 0.05 mag, according to the selection in distance (< 1 Kpc). A wrong estimation of the extinction would lead to a wrong estimation of the effective temperature, in the sense that systematically low values of A_V would give systematically low values of temperatures.

(ii) We assume a fixed metallicity (solar) both for the hot sd and the companion. If one (or both) component was a metal-poor object, the line-blanketing in the UV would be less pronounced which would translate into an increase of the flux at short wavelengths. The composite SED would show a flux excess in the GALEX range leading to a bad fitting to solar metallicity models in this regime. However, these objects would fall in the category *'binaries with bad SED fitting* and, thus, their temperatures would have not been estimated. Moreover, as we are selecting objects at less than 1 Kpc, the fraction of thick disc and halo stars is negligible. Therefore, we can discard metallicity as one of the main effects causing the spread in effective temperatures.

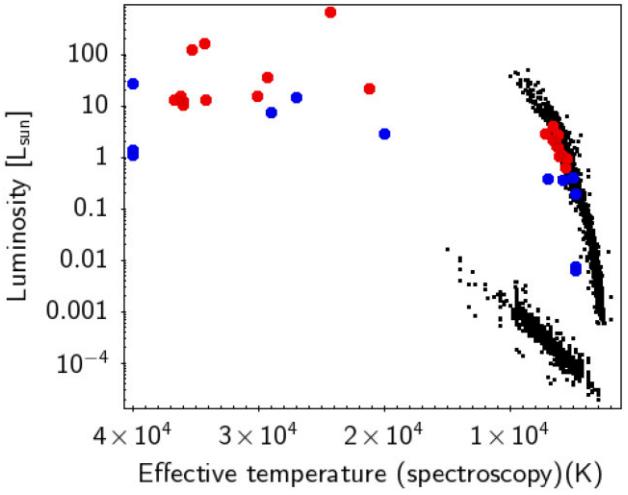

Figure 6. Position in an L–$T_{\rm eff}$ diagram of the Girven et al. (2012) (blue) and Németh et al. (2012) (red) objects analysed in Section 3.2. No subgiant/giant companions have been identified in our sample. Black dots are a subsample of the small black dots shown in Fig. 1 representing the main sequence and the white dwarf cooling track.

(iii) Girven et al. (2012) pointed out that the spectroscopic analysis of the composite hot sd is often carried out using a single-star model. However, the companion may have a significant contribution to the absorption lines used in the analysis resulting, thus, in a biased estimation of the system parameters, in particular of the spectroscopically derived effective temperatures.

(iv) In all the binary systems studied in this work, the companions have been assumed to be main-sequence stars. However, examples of giants and subgiants companions to hot sds have been previously reported (Heber et al. 2002). To discard this possibility, we have placed the hot sds and their companions in a luminosity–effective temperature diagram. Fig. 6 shows the position of the hot sds and their companions included in Fig. 5. We see how all our companions lie on or just below the main sequence, which discards the presence of subgiant/giant companions. We note that the range validity of the computed luminosity values is severely hampered by the small sample size (17 systems) available. Besides, the lack of photometric data bluer than GALEX data largely influences real uncertainties associated with temperature and, in turn, luminosity determinations. Other considerations, such as ageing or He content could set hot sds at higher (or lower) luminosity values (Kawka et al. 2015).

(v) As Rauch (2008) suggested, different physical approaches in the model atmosphere codes, in particular in the UV range, can translate into non-negligible differences in the values of the derived physical parameters. Girven et al. (2012) use Heber, Reid & Werner (2000) while Németh et al. (2012) rely on Hubeny & Lanz (1995) models. A detailed comparison between the outputs of these theoretical models – as well as the different methodologies used in Girven et al. (2012) and Németh et al. (2012) – goes beyond the scope of this paper.

Once we have confirmed that our methodology provides reliable values of $T_{\rm eff}$, we applied it to the 615 objects included in Geier (2020) and classified as binaries by VOSA. After applying the

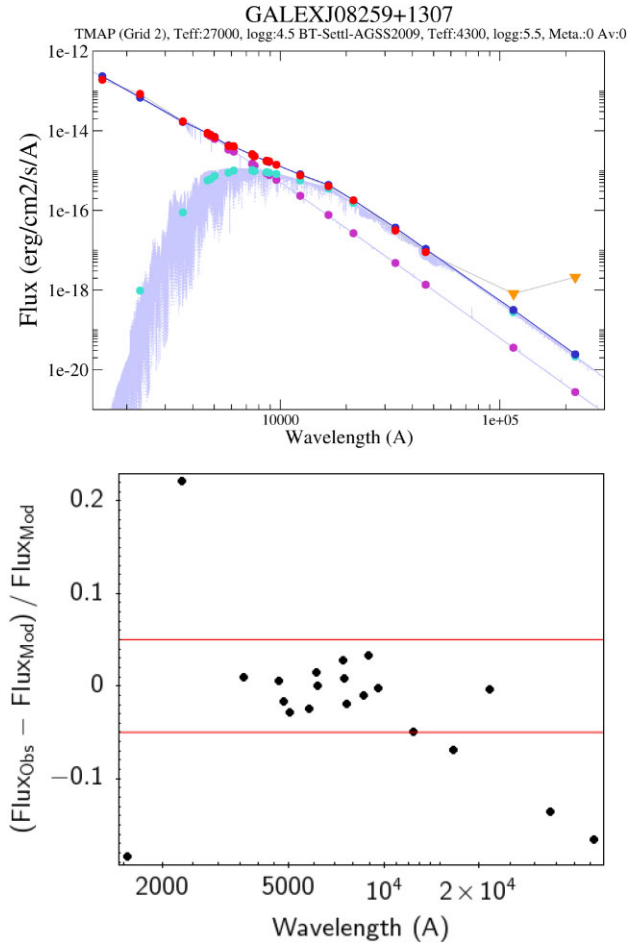

Figure 7. Top: Example of a two-body SED fitting using VOSA. The purple and green dots represent the TMAP and BT-Settl photometry of the model that fits best, respectively. The corresponding theoretical spectra are both plotted in purple. Red dots are the observed photometric points while the blue line and dots indicate the composite model that best fits the data. The yellow inverted triangles indicate that the photometric values correspond to upper limits and, thus, are not taken into account in the fitting process. Bottom: Residuals of the fit for the good photometric points (black dots). Red lines indicate a difference of 5 per cent between the observed and the model flux.

filtering criteria described in Section 3.1 (i.e. conditions in range of temperatures, distances, and UV photometry), only 70 sources were kept. These objects were visually inspected and only 42 of them showed a good SED two-body fitting (Fig. 7, as an example). The effective temperature distribution of the companions is shown in Fig. 8. We can see how the peak in the distribution is reached at $T_{\rm eff}$ = 5000–5500 K (G7V–K2V), while 75 per cent of the objects lie in the range 4000–6000 K (F9V–K8V). Particularly interesting are the three objects with substellar temperatures ($T_{\rm eff}$ < 2500 K), which would require further and more detailed analysis to confirm their substellar nature. Information on these 42 objects can be found in the SVO archive of hot sds (see Appendix).

4 CONCLUSIONS

A substantial fraction of hot sds reside in binaries (see Kawka et al. 2015; Pelisoli et al. 2020, and references therein). These are important objects to constrain the different evolution scenarios proposed to explain the loss of the hydrogen envelope during the

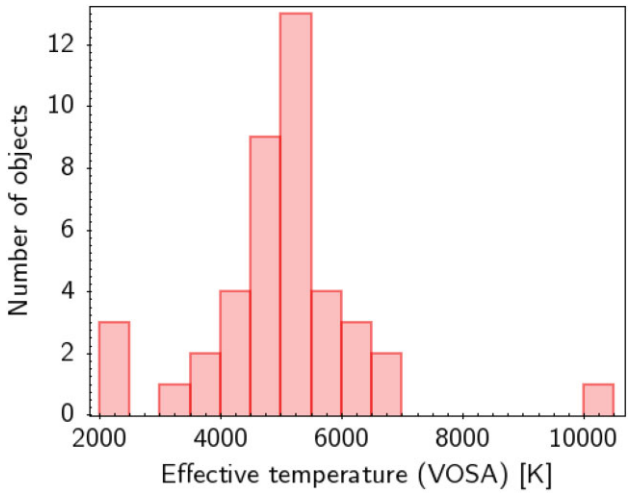

Figure 8. Effective temperatures of the cool companions of the 42 hot sds binaries for which physical parameters were derived from the SED two-body fitting. See Section 3.2 for details.

red giant branch phase. In this work, we have described a method to identify hot sd binaries with the help of the VOSA Virtual Observatory tool. The method consists on three steps: (i) Building of the SED from the UV to the infrared. (ii) Identification of binaries based on flux excess towards redder bands. (iii) Physical parameter estimation for both single and composite hot sds from the model that best fits.

Our procedure was checked against the hot sd catalogue of Geier (2020), the latest and most complete catalogue of hot sds presently available. From a sample of the 3186 objects with good photometric and astrometric information in *Gaia* EDR3 and lying on the expected locus for hot sds, VOSA identified 2469 single objects and 615 binary systems, showing a superior performance to the colour-based criterion proposed in Geier (2020) and finding 269 binaries not reported in that paper. The rest of objects (102) were not classified by VOSA due to their bad SED fitting. This gives a binary ratio of 20 per cent, a number that should be regarded as a lower limit for the binarity incidence which, in any case, lies within expectations. For the 2469 single hot sds, we computed temperatures, luminosities, and radii for 192 objects. Likewise, out of the 615 objects identified as binaries by VOSA, physical parameters could be reliably estimated for 42 of them using a two-body SED fitting.

According to Németh (2020), binary companions to hot sds could be distributed as follows: 13–17 per cent white dwarf, 26–30 per cent low-mass MS or near the substellar limit, and about 20 per cent more massive MS. Besides, MS + hot sd segregate into two main groups: (i) systems formed through common-envelope evolution would display orbital periods of less than 30 d and a low-mass companion; (ii) systems formed through Roche lobe overflow have periods in the range of 400–1500 d and more massive F–G–K companions. Ideally, a careful spectral decomposition analysis would also reveal reflection effects for the shortest period systems, chemical compositions, masses, ages, or rotational properties. In this context, our results provide an excellent bona fide hot sd binary candidates sample to be spectroscopically followed up by potentially interested researchers in the field.

Also, the application of Artificial Intelligence (AI) techniques to larger hot sd candidate data sets, such as, for example, the all-sky catalogue of 39 800 hot subluminous star candidates (Geier et al. 2019), and the use of new *Gaia* data releases could be considered for

future studies. Deep Learning techniques such as autoencoders and convolutional networks have proven useful in classifying objects with incomplete or noisy data, and can provide practical classifications if sufficient reference data are available to train the networks.

ACKNOWLEDGEMENTS

This work has been funded by RTI2018-095076-B-C22 and also by MCIN/AEI/10.13039/501100011033 through grant PID2020-112949GB-I00 and MDM-2017-0737 at Centro de Astrobiología (CSIC-INTA), Unidad de Excelencia María de Maeztu. We acknowledge support from CIGUS CITIC, funded by Xunta de Galicia and the European Union (FEDER Galicia 2014-2020 Program) through grant ED431G 2019/01; research consolidation grant ED431B 2021/36; and scholarship from Xunta de Galicia 620 and the European Union (European Social Fund - ESF) ED481A-2019/155. AA acknowledges support from Government of Comunidad Autónoma de Madrid (Spain) through postdoctoral grant 'Atracción de Talento Investigador' 2018-T2/TIC-11697. This publication makes use of VOSA, developed under the Spanish Virtual Observatory project. This research has made use of Aladin sky atlas developed at CDS, Strasbourg Observatory, France (Bonnarel et al. 2000; Boch & Fernique 2014). Vizier (Ochsenbein, Bauer & Marcout 2000), TOPCAT (Taylor 2005), and STILTS (Taylor 2006) have also been widely used in this paper.

DATA AVAILABILITY

The data underlying this article are available in the SVO archive of hot subdwarfs at http://svocats.cab.inta-csic.es/hsa2/.

REFERENCES

Alam S. et al., 2015, ApJS, 219, 12
Allard F., Homeier D., Freytag B., 2012, Phil. Trans. R. Soc. A, 370, 2765
Arenou F. et al., 2018, A&A, 616, A17
Bayo A., Rodrigo C., Barrado Y Navascués D., Solano E., Gutiérrez R., Morales-Calderón M., Allard F., 2008, A&A, 492, 277
Bianchi L., GALEX Team, 2000, Mem. Soc. Astron. Ital., 71, 1123
Boch T., Fernique P., 2014, in Manset N., Forshay P., eds, ASP Conf. Ser. Vol. 485, Astronomical Data Analysis Software and Systems XXIII. Astron. Soc. Pac., San Francisco, p. 277
Bonnarel F. et al., 2000, A&AS, 143, 33
Camarota L., Holberg J. B., 2014, MNRAS, 438, 3111
Chambers K. C. et al., 2016, preprint (arXiv:1612.05560)
D'Cruz N. L., Dorman B., Rood R. T., O'Connell R. W., 1996, ApJ, 466, 359
Drilling J. S., Jeffery C. S., Heber U., Moehler S., Napiwotzki R., 2013, A&A, 551, A31
Fitzpatrick E. L., 1999, PASP, 111, 63
Gaia Collaboration, 2020, VizieR Online Data Catalog, p. I/350
Geier S. et al., 2011, A&A, 530, A28
Geier S. et al., 2015, A&A, 577, A26
Geier S., 2020, A&A, 635, A193
Geier S., Østensen R. H., Nemeth P., Gentile Fusillo N. P., Gänsicke B. T., Telting J. H., Green E. M., Schaffenroth J., 2017, A&A, 600, A50
Geier S., Raddi R., Gentile Fusillo N. P., Marsh T. R., 2019, A&A, 621, A38
Girven J. et al., 2012, MNRAS, 425, 1013
Heber U., 2009, ARA&A, 47, 211
Heber U., 2016, PASP, 128, 082001
Heber U., Moehler S., Napiwotzki R., Thejll P., Green E. M., 2002, A&A, 383, 938
Heber U., Reid I. N., Werner K., 2000, A&A, 363, 198
Henden A. A., 2015, AAS, id.336.16
Hubeny I., Lanz T., 1995, ApJ, 439, 875
Indebetouw R. et al., 2005, ApJ, 619, 931
Kawka A., Vennes S., O'Toole S., Németh P., Burton D., Kotze E., Buckley D. A. H., 2015, MNRAS, 450, 3514
Kepler S. O. et al., 2015, MNRAS, 446, 4078
Kepler S. O. et al., 2016, MNRAS, 455, 3413
Kilkenny D., Heber U., Drilling J. S., 1988, South Afr. Astron. Obs. Circ., 12, 1
Lada C. J. et al., 2006, AJ, 131, 1574
Lei Z., Zhao J., Németh P., Zhao G., 2019, ApJ, 881, 135
Lindegren L. et al., 2018, A&A, 616, A2
Lindegren L. et al., 2021, A&A, 649, A2
Luri X. et al., 2018, A&A, 616, A9
Magnier E. A. et al., 2020, ApJS, 251, 6
Németh P., 2020, Contrib. Astron. Obs. Skalnate Pleso, 50, 546
Németh P., Kawka A., Vennes S., 2012, MNRAS, 427, 2180
Ochsenbein F., Bauer P., Marcout J., 2000, A&AS, 143, 23
Oreiro R., Rodríguez-López C., Solano E., Ulla A., Østensen R., García-Torres M., 2011, A&A, 530, A2
Pelisoli I., Vos J., Geier S., Schaffenroth V., Baran A. S., 2020, A&A, 642, A180
Pérez-Fernández E., Ulla A., Solano E., Oreiro R., Rodrigo C., 2016, MNRAS, 457, 3396
Rauch T., 2008, A&A, 481, 807
Rauch T., Deetjen J. L., 2003, in Hubeny I., Mihalas D., Werner K., eds, ASP Conf. Ser. Vol. 288, Stellar Atmosphere Modeling. Astron. Soc. Pac., San Francisco, p. 103
Riello M. et al., 2021, A&A, 649, A3
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, ASP Conf. Ser. Vol. 347, Astronomical Data Analysis Software and Systems XIV. Astron. Soc. Pac., San Francisco, p. 29
Taylor M. B., 2006, in Gabriel C., Arviset C., Ponz D., Enrique S., eds, ASP Conf. Ser. Vol. 351, Astronomical Data Analysis Software and Systems XV. Astron. Soc. Pac., San Francisco, p. 666
Wolf C. et al., 2018, Publ. Astron. Soc. Aust., 35, e010
Wright E. L. et al., 2010, AJ, 140, 1868

APPENDIX A: VIRTUAL OBSERVATORY COMPLIANT, ONLINE CATALOGUE

In order to help the astronomical community on using our catalogue of hot sds candidates, we developed an archive system that can be accessed from a webpage[8] or through a Virtual Observatory ConeSearch.[9]

The archive system implements a very simple search interface that allows queries by coordinates and radius as well as by other parameters of interest. The user can also select the maximum number of sources (with values from 10 to unlimited). The result of the query is an HTML table with all the sources found in the archive fulfilling the search criteria. The result can also be downloaded as a VOTable or a CSV file. Detailed information on the output fields can be obtained by placing the mouse over the question mark located close to the name of the column. The archive also implements the SAMP[10] (Simple Application Messaging) Virtual Observatory protocol. SAMP allows Virtual Observatory applications to communicate with each other in a seamless and transparent manner for the user. This way, the results of a query can be easily transferred to other VO applications, such as, for instance, Topcat.

[8] http://svocats.cab.inta-csic.es/hsa2/
[9] e.g. http://svocats.cab.inta-csic.es/hsa2/cs.php?RA=155.841&DEC=-37.617&SR=0.1&VERB = 2.
[10] http://www.ivoa.net/documents/SAMP

This paper has been typeset from a TEX/LATEX file prepared by the author.